昆虫学报 *Acta Entomologica Sinica*, February 2017, 60(2): 136 – 147　　doi:10.16380/j.kcxb.2017.02.002

中华按蚊化学感受蛋白(CSP)家族基因的全基因组鉴定和特征分析

梅　婷, 何正波, 王小婷, 王婷婷, 陈　斌*

(重庆师范大学生命科学学院, 昆虫与分子生物学研究所, 媒介昆虫重庆市重点实验室, 重庆 401331)

摘要:【目的】在全基因组水平鉴定中华按蚊 *Anopheles sinensis* 化学感受蛋白(chemosensory protein, CSP)家族基因,预测该家族基因的特征,研究代表性双翅目昆虫 CSP 基因的系统发育和进化。【方法】在 NCBI 数据库中下载 CSP 氨基酸序列作为询问序列,通过 Blast 和 HMM 方法在全基因组水平搜索和鉴定中华按蚊、冈比亚按蚊 *An. gambiae*、埃及伊蚊 *Aedes aegypti* 和致倦库蚊 *Culex quinquefasciatus* 的 CSP 家族基因并命名;通过生物信息学方法预测中华按蚊 CSP 家族基因的特性(基因结构、基因组定位、剪切模式、Ka/Ks 比值)、保守结构域和蛋白质结构等;通过 MEGA 软件用最大似然法(maximum likelihood, ML)推断该家族基因的系统发育。【结果】中华按蚊、冈比亚按蚊、埃及伊蚊和致倦库蚊基因组分别有 8, 8, 43 和 27 个 CSP 家族基因。中华按蚊 CSP 家族基因(*AsCSPs*)都有全长的转录组序列,编码 116(*AsCSP*7) ~ 335(*AsCSP*5)个氨基酸;7 个 *AsCSPs* 分布于 Scaffold51 上,*AsCSP*8 分布于 Scaffold116;*AsCSP*1 ~ *AsCSP*8 分别有 3, 2, 1, 1, 1, 2, 1 和 2 个选择性剪切模式;*AsCSP*3 的表达量最高,其 FPKM 值达到 385.46。AsCSPs 的 N 端信号肽由 17 ~ 37 个氨基酸组成,均含有 4 个保守的半胱氨酸位点(CYS68, CYS75, CYS94 和 CYS97),这些位点界定了两个二硫键(CYS68-CYS75 和 CYS94-CYS97)。系统发育分析结果显示,4 种蚊虫的 8 个 CSP 基因各自形成明显的支系,被命名为组 CSP1 ~ CSP8 (CSP1 – CSP8 group)。埃及伊蚊和致倦库蚊分别有 35 和 18 个 CSP 基因形成了一个不为按蚊所共有的特殊支系,被命名为 Culicinae-specific 组。替换率结果显示,中华按蚊与冈比亚按蚊同源基因对的 Ka/Ks 值都小于 1,说明 CSP 家族基因在进化过程中主要是受到纯化选择作用。【结论】本研究为蚊虫,特别是中华按蚊基因组 CSP 家族基因提供了信息框架,为进一步开展该家族基因的功能研究奠定了基础。

关键词: 中华按蚊; 化学感受蛋白; 全基因组鉴定; 生物信息学; 基因特征; 系统发育

中图分类号: Q966　　**文献标识码:** A　　**文章编号:** 0454-6296(2017)02-0136-12

Genome-wide identification and characterization of genes of the chemosensory protein (CSP) family in *Anopheles sinensis* (Diptera: Culicidae)

MEI Ting, HE Zheng-Bo, WANG Xiao-Ting, WANG Ting-Ting, CHEN Bin* (Chongqing Key Laboratory of Vector Insects, Institute of Entomology and Molecular Biology, College of Life Sciences, Chongqing Normal University, Chongqing 401331, China)

Abstract:【Aim】The study aims to identify the genes of the chemosensory protein (CSP) family in *Anopheles sinensis* whole-genome, to predict the characteristics of these CSP genes, and to investigate the phylogenetics and evolution of the CSPs in representative dipteran species. 【Methods】We searched, identified and named the CSP genes in *An. sinensis*, *An. gambiae*, *Aedes aegypti* and *Culex quinquefasciatus* genomes with Blast and HMM methods using CSP amino acid sequences downloaded from

基金项目: "两江学者"计划专项经费; 国家自然科学基金项目(31672363, 31372265); 国际原子能机构 CRP 项目(18268); 国家科技基础性工作专项重点项目(2015FY210300)

作者简介: 梅婷, 女, 1990 年 7 月生, 重庆开州区人, 硕士研究生, 研究方向为昆虫分子生物学, E-mail: 393312720@qq.com

* 通讯作者 Corresponding author, E-mail: bin.chen@cqnu.edu.cn

收稿日期 Received: 2016-10-18; 接受日期 Accepted: 2016-12-21

NCBI database as inquiry entry, analyzed the characteristics of the CSP genes in *An. sinensis* using bioinformatics methods, including the structure, location, splicing and Ka/Ks ratio of these genes, conservative domains and protein structures, and deduced the phylogeny of CSP genes using maximum likelihood (ML) method with MEGA software. 【Results】The genomes of *An. sinensis*, *An. gambiae*, *Ae. aegypti* and *Cx. quinquefasciatus* contain 8, 8, 43 and 27 CSP genes, respectively. The CSP genes identified in *An. sinensis* (*AsCSPs*) are all supported by full-length transcripts, encoding 116 (*AsCSP*7) to 335 (*AsCSP*5) amino acids. Out of them, seven *AsCSPs* are located on Scaffold 51 and *AsCSP*8 on Scaffold116. *AsCSP*1 - *AsCSP*8 possess 3, 2, 1, 1, 1, 2, 1 and 2 splicing variants, respectively, and *AsCSP*3 shows the highest expression level with a FPKM value of 385.46. All AsCSPs each has a N-terminal signal peptide consisting of 17 - 37 amino acids, and contains 4 conserved cysteine sites (CYS68, CYS75, CYS94 and CYS97), which define two disulfide bonds (CYS68-CYS75 and CYS94-CYS97). Phylogenetic analysis results showed that eight CSP genes in four mosquito species are each grouped into a significant clade, and were named as CSP1 - CSP8 group, respectively. Thirty-five and 18 CSP genes in *Ae. aegypti* and *Cx. quinquefasciatus*, respectively, are grouped into a special clade without homologous genes with *Anopheles* species, and was named as the Culicinae-specific group in this study. The Ka/Ks values of orthologous gene pairs between *An. sinensis* and *An. gambiae* were all less than 1, suggesting that the CSP gene family mainly experienced purifying selection during evolution. 【Conclusion】The study provides an information frame of the CSP gene family in mosquito species, especially *An. sinensis*, and lays the foundation for further functional analysis of these genes.

Key words: *Anopheles sinensis*; chemosensory protein; genome-wide identification; bioinformatics; gene characteristics; phylogeny

昆虫在长期进化过程中形成了一套感知外界环境各种信号物质刺激的嗅觉感受系统,通过这套感受系统,昆虫可识别来自种内、种间和外界环境的各种复杂化学信息,再将这些信息转换成体内的电子信号传达到中枢神经系统,从而使昆虫判断出食物、天敌和同伴并引起取食、趋避、栖境选择及"社会"交往、求偶行为等一系列行为反应(Sanchez-Gracia *et al*., 2009; Pelletier and Leal, 2011; Gu *et al*., 2012)。昆虫的化学感受器多为毛发状空心结构,内部的水溶状淋巴液中浸润着双向感觉神经元,还包括 2 类水溶性蛋白:气味结合蛋白(odorant binding protein, OBP)和化学感受蛋白(chemosensory protein, CSP)(Monteforti *et al*., 2002; 秦赠等, 2014)。与 OBPs 不同的是,CSP 家族基因主要结合和运载非挥发性的气味分子和化学刺激物(Ban *et al*., 2002)。从 CSP 的研究历程来看,其命名并未统一。起初因其在触角中优先表达被称为嗅觉特异-D 蛋白质(olfactory-specific-D proteins, OS-D)或信息素结合蛋白 A-10(pheromone-binding proteins A-10)(McKenna *et al*., 1994; Pikielny *et al*., 1994),后又发现该家族基因在一些昆虫中的感觉器官中(如下唇须和足等)也有表达,将其称为感受器官蛋白(sensory appendage protein, SAP)(Robertson *et al*., 1999; Angeli *et al*., 1999)。实际上,它们具有相同的半胱氨酸保守域(cysteine pattern, "CSP" motif),属于同一类基因,是穿梭于嗅觉神经树突末端的水溶性淋巴液中的一类载体蛋白质(Forêt *et al*., 2007),可作为气味分子和化学信号刺激物的载体蛋白,并将其转运到感受器树突膜上的受体位点,使之与受体蛋白结合。因此,当前这类蛋白统称为化学感受蛋白(CSP)(Schneider, 1969; Pelosi *et al*., 2005, 2006; Pelletier and Leal, 2011; Iovinella *et al*., 2013)。

自 Pikielny 等(1994)在果蝇 *Drosophila* 的触角中首次发现 CSP 家族基因以来,CSP 被发现存在于各类昆虫中,其基因数量在各物种中差异很大,在黑腹果蝇 *D. melanogaster* 中仅有 4 个,而在埃及伊蚊中达到 43 个(Pelletier and Leal, 2011; Vieira and Rozas, 2011; Andersson *et al*., 2013; Iovinella *et al*., 2013; Dippel *et al*., 2014; Wang *et al*., 2015)。研究发现 CSP 类蛋白是一类相对分子质量都较小的可溶性结合蛋白质,具有 4 个保守的半胱氨酸位点,构成两个二硫键(C1-C2 和 C3-C4),具有 6 个 α 螺旋(Briand *et al*., 2002; Li *et al*., 2003; Wanner *et al*., 2004)。早期的研究指出,CSP 主要是在嗅觉器官中参与结合转运外界气味分子和化学信号刺激物

(Jacquin-Joly *et al*., 2001; Lartigue *et al*., 2002; Ozaki *et al*., 2005),但随后也有大量研究表明 CSP 在一些非感官组织中有表达,反映出 CSPs 在调节昆虫生长发育(Wanner *et al*., 2005)、胚胎发育(Forêt *et al*., 2007; Maleszka *et al*., 2007)、生理节律(Ozaki *et al*., 2005)及免疫系统调节(Sabatier *et al*., 2003; Jin *et al*., 2005)等方面有重要的生理功能。

随着新一代测序技术的广泛应用,越来越多的物种全基因组测序工作已完成,使得一些物种的 CSP 家族基因被陆续鉴定出来,包括西方蜜蜂 *Apis mellifera*、家蚕 *Bombyx mori*、蚂蚁 *Camponotus japonicus*、中华通草蛉 *Chrysoperla sinica*、红脂大小蠹 *Dendroctonus valens* 等昆虫(Forêt *et al*., 2007; Gong *et al*., 2007; Hojo *et al*., 2015; Li *et al*., 2015; Gu *et al*., 2015)。中华按蚊 *Anopheles sinensis* 是中国记述最早、研究最多的蚊虫,是我国疟疾的主要传播媒介,也是马来丝虫病的重要媒介之一,在有些地区也曾从中华按蚊中分离到流行性乙型脑炎病毒(Chow, 1991; Zhang *et al*., 1994; Manguin *et al*., 2010; Chang *et al*., 2014)。近年来我们实验室开展了中华按蚊基因组和转录组精细测序和注释,中华按蚊 CSP 家族基因尚未有研究报道。本研究鉴定了中华按蚊全基因组上的 CSP 家族基因,分析了这些基因的序列和功能特征,在基因组上的定位、结构和剪切模式,系统发育关系及非同义替换等,目的在于提供该家族基因的基本信息框架,为进一步开展外界信息素的感受机制研究奠定基础。

1 材料与方法

1.1 基因序列数据库

中华按蚊基因组数据和转录组数据(SRA 登录号: SRA073189)来自重庆师范大学昆虫与分子生物学研究(Chen *et al*., 2014),其他物种包括冈比亚按蚊 *Anopheles gambiae*、埃及伊蚊 *Aedes aegypti*、致倦库蚊 *Culex quinquefasciatus* 和黑腹果蝇的基因组数据均来自于 VectorBase 数据库(https://www.vectorbase.org/)。

1.2 中华按蚊 CSP 家族基因在全基因组上的鉴定和特征分析

基于发表的家蚕、意大利蜜蜂 *Apis mellifera ligustica* 和赤拟谷盗 *Tribolium castaneum* 等的 CSP 氨基酸序列(Pelletier *et al*., 2011; Vieira and Rozas, 2011)在 NCBI 数据库中下载 CSP 家族基因的同源序列,将这些氨基酸序列作为询问序列,首先用 BLASTP 和 TBLASTN 搜索中华按蚊氨基酸和基因组数据库($E \leqslant 10^{-5}$);同时采用 HMM Pfam 号(pfam03392)搜索获得一个新的 CSP 家族基因组序列。将所有获得的基因组 CSP 序列合并去除重复,并通过 FGENESH +(http://linux1.softberry.com/)在线软件对提取出的基因组序列进行预测,然后将预测出的氨基酸序列与 NCBI 的 GenBank 数据库进行比对,并采用 SMART(http://smart.embl-heidelberg.de/)软件分析是否存在 OS-D 结构域,根据其结构域信息初步确定是否为 CSP 候选基因。基于 VectorBase 和 NCBI 数据库的基因组数据,用同样方法鉴定几种代表性双翅目昆虫 CSP 家族基因。

将以上得到的候选 CSP 家族基因通过 ProtParam(http://www.expasy.org/protparam/)、SignalP(http://www.cbs.dtu.dk/services/SignalP/)、TargetP(http://www.cbs.dtu.dk/services/TargetP/)、TMHMM(http://www.cbs.dtu.dk/services/TMHMM/)软件对其蛋白序列分子式、分子量、等电点、信号肽、亚细胞定位和跨膜区等进行预测分析。

1.3 中华按蚊 CSP 家族基因在基因组上的定位

为了定位中华按蚊 CSP 家族基因(*AsCSPs*)在全基因组上的位置,我们将中华按蚊候选 CSP 家族基因的氨基酸序列通过 TBLASTN 搜索中华按蚊基因组数据库,统计各基因在基因组上 Scaffold 的具体位置。冈比亚按蚊 CSP 家族基因(*AgCSPs*)在基因组上的位置数据下载于 VectorBase 数据库,其中 *AgCSP*6 的定位参考 Iovinella 等(2013)(基因组数据为 Anopheles-gambiae-PEST_SCAFFOLDS_AgamP4)。用 Photoshop CS8 绘制 AsCSPs 和 AgCSPs 基因在基因组上的位置图。

1.4 中华按蚊 CSP 家族基因的转录、表达和结构分析

使用中华按蚊 CSP 候选基因序列搜索中华按蚊转录组数据库($E \leqslant 10^{-5}$),检测 CSP 基因是否存在转录本,并获得 CSP 基因家族成员对应的 FPKM(fragment per kb per million reads)值,以此估测其表达量。

通过 Softberry 软件和中华按蚊 CSP 家族基因的转录组序列预测中华按蚊 CSP 家族基因的结构,通过在线软件 GSDS(Gene Structure Display Server)(http://gsds.cbi.pku.edu.cn/index.php)预测各基因内含子相位(Hu *et al*., 2015),使用 Photoshop CS3 软件绘制各基因的外显子-内含子结构。结合结构

预测，将中华按蚊 CSP 家族基因的转录组序列通过 BLASTN 反比到对应基因组序列上，确定中华按蚊 CSP 家族基因选择性的剪切模式。

1.5　中华按蚊 CSP 家族基因的保守结构域和蛋白质结构分析

使用 ClustalW 软件对中华按蚊 CSP 家族氨基酸序列进行多重序列比对，将其保存为 Fasta 格式文件，再导入 GeneDoc 软件预测其保守区域。利用在线软件 PSIPREDv3.3（http://bioinf.cs.ucl.ac.uk/psipred/）预测所有候选 CSP 家族基因的蛋白质二级结构；采用 SWISS-MODEL（http://swissmodel.expasy.org/）进行同源建模及 3D 结构预测分析（唐尧等，2014；刘柏琦等，2016）。

1.6　双翅目代表性种 CSP 家族基因的系统发育关系分析

通过软件 ClustalW 进行氨基酸序列多重比对，以黑腹果蝇 CSP 氨基酸序列为外群，采用 MEGA5.0 软件使用最大似然法（maximum likelihood，ML）（Partial deletion 处理缺失数据，WAG + G 模型）对中华按蚊、冈比亚按蚊、埃及伊蚊和致倦库蚊构建系统发育树，使用 1 000 次重复计算系统发育树上的 bootstrap 值，将大于 50% 的 bootstrap 值标于系统发育树上，讨论其系统发育关系。

1.7　中华按蚊 CSP 家族基因的 Ka/Ks 分析

将中华按蚊和冈比亚按蚊的 CSP 家族基因的蛋白质编码序列（CDs）的终止密码子去除，采用 MEGA5.0 对每个中华按蚊 CSP 家族 CD 序列与冈比亚按蚊的直系同源基因 CD 序列分别进行比对，去掉 gap 后，再将其结果用 PAML 软件（Yang and Nielsen，2000）计算非同义替换率（Ka）、同义替换率（Ks）及它们的比值（Ka/Ks），分析中华按蚊 CSP 家族基因在进化历史中的选择压力和选择模式。

2　结果

2.1　中华按蚊 CSP 家族基因在全基因组上的鉴定和特征分析

通过在中华按蚊全基因组上的搜索与鉴定，获得了中华按蚊 CSP 家族 8 个候选基因的全长序列。根据它们的 OS-D 结构域信息及其在基因组上的位置，分别将其命名为 *AsCSP*1，*AsCSP*2，*AsCSP*3，*AsCSP*4，*AsCSP*5，*AsCSP*6，*AsCSP*7 和 *AsCSP*8。这 8 个基因的氨基酸序列与冈比亚按蚊同源基因的序列一致性达到 67% ~ 95%，氨基酸数目为 116（*AsCSP*7）~ 335（*AsCSP*5）个，等电点为 4.71 ~ 9.76，分子量为 12.9 ~ 36.0 kD（表 1）。

表 1　中华按蚊 CSP 家族基因特征及其与冈比亚按蚊 CSP 基因的比较

Table 1　Characteristics of CSP genes from *Anopheles sinensis* and the comparison with those *in An. gambiae*

基因名称 Gene name	基因组上的定位 Location on scaffold	基因全长 Gene size (bp)	氨基酸大小 Amino acid size (a. a.)	分子量 Molecular weight (kD)	与冈比亚按蚊同源基因比对 Comparison with homologous gene of *An. gambiae*	
					基因 ID 号 Gene ID	一致性（%） Identity
*AsCSP*1	Scaffold51：4 139 817 – 4 140 197	381	126	14.7	AAL84186	83
*AsCSP*2	Scaffold51：4 143 486 – 4 143 869	384	127	14.7	XP_317407	95
*AsCSP*3	Scaffold51：4 155 411 – 4 155 791	381	126	14.6	XP_317405	91
*AsCSP*4	Scaffold51：4 158 635 – 4 159 239	605	174	19.5	XP_317404	67
*AsCSP*5	Scaffold51：4 182 389 – 4 197 875	15 487	335	36	XP_317401	85
*AsCSP*6	Scaffold51：4 213 890 – 4 215 154	1 265	123	14.2	CAG26923	90
*AsCSP*7	Scaffold51：4 238 749 – 4 259 876	21 128	116	12.9	XP_317397	94
*AsCSP*8	Scaffold116：3 632 275 – 3 637 193	4 919	129	14.7	AGF68546	81

基因名称 Gene name	分子式 Molecular formula	等电点（pI） Isoelectric point	跨膜区 Transmembrane domain	亲水性系数 Hydrophobicity coefficient	不稳定指数 Instability index
*AsCSP*1	$C_{564}H_{1041}N_{177}O_{196}S_5$	8.29	–	–0.587	34.34
*AsCSP*2	$C_{658}H_{1048}N_{172}O_{199}S_6$	5.34	–	–0.554	34.77
*AsCSP*3	$C_{645}H_{1027}N_{175}O_{198}S_6$	8.26	–	–0.717	37.65
*AsCSP*4	$C_{838}H_{1329}N_{231}O_{279}S_{11}$	4.71	–	–0.673	45.94
*AsCSP*5	$C_{1571}H_{2508}N_{458}O_{501}S_6$	5.71	–	–0.364	43.41
*AsCSP*6	$C_{627}H_{1042}N_{176}O_{188}S_6$	9.35	–	–0.698	25.48
*AsCSP*7	$C_{564}H_{937}N_{165}O_{160}S_9$	9.76	+	0.059	47.08
*AsCSP*8	$C_{609}H_{999}N_{173}O_{182}S_{10}$	9.17	+	–0.042	50.65

+：有 Present；–：没有 Absent.

对这 8 个 AsCSPs 的氨基酸序列进行多重比对和特征分析，结果表明，这些序列均含有 4 个保守的半胱氨酸位点（CYS68，CYS75，CYS94 和 CYS97），构成两个二硫键（CYS68-CYS75 和 CYS94-CYS97），这与前期研究的化学感受蛋白家族基因的标志性特征一致（图 1）；N 端含有 17～37 个氨基酸组成的信号肽；亚细胞定位也表明这 8 个基因的蛋白都属于分泌途径蛋白（secretory pathway signal peptide），可能在信号物质的识别和传导中起重要作用（Vieira and Rozas，2011）。AsCSP7 和 AsCSP8 的 OS-D 保守结构域具有 85 个氨基酸，AsCSP1－6 的 OS-D 保守结构域长 93 个氨基酸，序列比对显示这个区域高度保守和一致。所有这些化学感受蛋白中，只有 AsCSP7 和 AsCSP8 具有跨膜区。亲水性指数分析表明，除 AsCSP7 外，其他 7 个蛋白均为亲水性蛋白，推测 AsCSPs 家族可能是一个可溶性嗅觉蛋白家族（表 1，图 1）。

2.2 中华按蚊 CSP 家族基因在基因组上的定位

将中华按蚊 CSP 家族基因定位到基因组上，结果表明 *AsCSP*1－*AsCSP*7 均分布于 scaffold51 上，*AsCSP*8 分布于 Scaffold116 上。鉴定获得了 8 个冈比亚按蚊 CSP 家族基因，将这些基因定位到冈比亚按蚊基因组上，发现 *AgCSP*1－*AgCSP*7 均分布于染色体 3R 上，*AgCSP*8 分布于 scf_1106392397089（其染色体位置尚未定）上（图 2）。用 Photoshop CS8 软件绘制中华按蚊和冈比亚按蚊 CSP 家族基因在基因组上的相对位置并作比较，结果显示，除了 *CSP*8 以外，其余 7 个 CSP 家族基因均成簇分布于同一 Scaffold 或染色体上，并形成一一对应关系，仅 CSP1 和 CSP4 两个基因的序列方向不一致（图 2）。

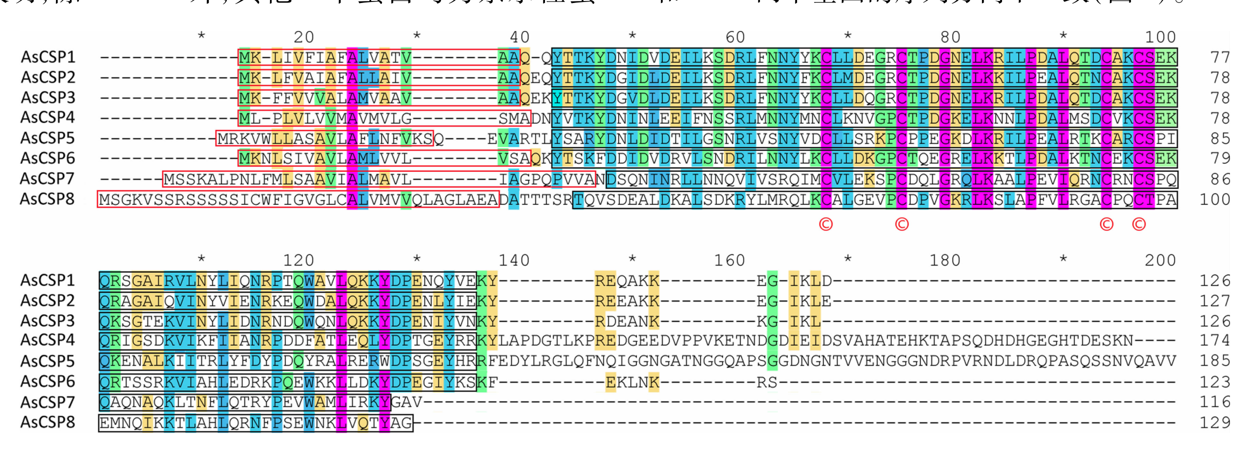


图 1 中华按蚊 CSP 家族基因氨基酸序列的多重比对和特征

Fig. 1 Multiple sequence alignment and characteristics of amino acid sequences of CSP family in *Anopheles sinensis*

AsCSP1－AsCSP8：中华按蚊化学感受蛋白家族基因 Genes of CSP family in *An. sinensis*. 下同 The same below. 各基因编码氨基酸序列 N 端信号肽序列标示在红色框内，OS-D 保守结构域标示在黑色框内；100%，80%～99%，60%～79% 和 <60% 一致性的氨基酸序列分别用紫色、蓝色、绿色和橙色标示；4 个保守的半胱氨酸位点用 *CC* 标示在序列下方；AsCSP5 共 335 个氨基酸，仅显示到第 185 位氨基酸处。The N-terminal signal peptides of each amino acid sequence are boxed in red, the OS-D conserved regions are boxed in black; purple, blue, green and orange shades denote the amino acid sequences with the identity of 100%, 80%－99%, 60%－79% and below 60%, respectively; the four characteristic cysteine sites are labeled with *CC* below the sequence; AsCSP5 is shown up to the 185th amino acid of 335 amino acids.

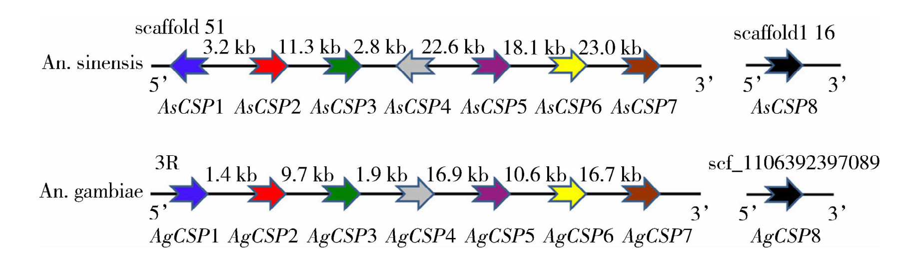


图 2 中华按蚊和冈比亚按蚊 CSP 基因在基因组上的定位

Fig. 2 Relative locations of the CSP genes in *Anopheles sinensis* and *An. gambiae* genomes

*AgCSP*1－*AgCSP*8：冈比亚按蚊化学感受蛋白家族基因 Genes of CSP family in *An. gambiae*. 不同的基因用不同颜色的箭头标示，箭头标示基因序列 5′－3′方向，基因间的距离标示在基因组序列的横线上方。Different genes are marked with different colors of arrows with the arrow direction denoting 5′－3′ of the gene sequence, and the distance between genes are marked above the line demonstrating genomic sequence.

2.3　中华按蚊 CSP 家族基因的转录、表达和结构

对中华按蚊的转录组数据库搜索表明，这 8 个 CSP 家族基因均具有转录本（表 2）。*AsCSP*3，*AsCSP*4，*AsCSP*5 和 *AsCSP*7 具有 1 个转录本，*AsCSP*2，*AsCSP*6 和 *AsCSP*8 具有两个选择性转录本，*AsCSP*1 具有 3 个选择性转录本。以 FPKM 值估测每个转录本的表达量，结果显示，Unigene871_5（*AsCSP*3）的表达丰度最高，FPKM 值达 385.46，其次是 Unigene27178_5（*AsCSP*2），CL1318. Contig3_5（*AsCSP*1）和 CL1733. Contig1_5（*AsCSP*6），其 FPKM 值分别为 71.60，10.25 和 9.15；CL1318. Contig1_5（*AsCSP*2）的表达丰度最低，FPKM 值仅为 0.36。总和各基因不同转录本的 FPKM 值，*AsCSP*3 的表达丰度最高（FPKM 值为 385.46），其次为 *AsCSP*2（71.96），*AsCSP*1（17.13），*AsCSP*6（12.68），*AsCSP*4（6.23），*AsCSP*7（5.98），*AsCSP*8（4.08）和 *AsCSP*5（2.73）。

表 2　中华按蚊 CSP 家族基因的选择性剪切及转录丰度

Table 2　Splicing variants and transcription richness of genes of the CSP family in *Anopheles sinensis*

基因名称 Gene name	选择性剪切 ID ID of splicing variants	表达丰度 FPKM	表达丰度总和 Total FPKM	外显子数量 Exon number
*AsCSP*1	CL1318. Contig2_5	5.46	17.13	1
	CL1318. Contig3_5	10.25		
	CL1318. Contig4_5	1.42		
*AsCSP*2	Unigene27178_5	71.60	71.96	1
	CL1318. Contig1_5	0.36		
*AsCSP*3	Unigene871_5	385.46	385.46	1
*AsCSP*4	Unigene21154_5	6.23	6.23	2
*AsCSP*5	CL1408. Contig4_5	2.73	2.73	5
*AsCSP*6	CL1733. Contig1_5	9.15	12.68	2
	CL1733. Contig2_5	3.53		
*AsCSP*7	Unigene16775_5	5.98	5.98	2
*AsCSP*8	CL2094. Contig1_5	3.27	4.077	3
	CL2094. Contig2_5	0.81		

表达丰度 FPKM：每百万个映射的读取序列中每千个碱基的外显子的序列数 Number of fragments per kilobase of exon model per million mapped reads.

基于基因组和转录组序列，通过在线软件 GSDS 绘制了 AsCSPs 基因结构图（图 3），结果显示这 8 个基因的外显子和内含子的分布模式差异较大，*AsCSP*1，*AsCSP*2 和 *AsCSP*3 仅有一个外显子，*AsCSP*4，*AsCSP*6 和 *AsCSP*7 有两个外显子，*AsCSP*5 和 *AsCSP*8 分别有 5 个和 3 个外显子（表 2）。仅 1 个内含子相位为 2 相位，其余都为 1 相位，*AsCSP*5，*AsCSP*6，*AsCSP*7 和 *AsCSP*8 内含子区长于蛋白质编码区，特别 *AsCSP*7 的唯一内含子长度达 20 782 bp。分析中华按蚊 CSP 家族基因的转录本发现，*AsCSP*1，*AsCSP*6 和 *AsCSP*8 的选择性剪切由 5′端不同长度的 UTR 所造成，*AsCSP*2 的选择性剪切由 3′端不同长度的 UTR 所造成，*AsCSP*4 和 *AsCSP*7 的转录本在 5′端分别缺失 160 和 134 bp 的氨基酸编码序列（图 3）。

2.4　中华按蚊 CSP 家族基因的蛋白质结构

根据在线软件 PSIPRED 和 SWISS-MODEL 预测了 AsCSPs 蛋白的空间结构，结果表明，除 AsCSP5 外的 7 个 AsCSPs 的二级结构都是由 6 个 α 螺旋结构和直链串联组成（图 4：A），这是昆虫气味结合蛋白二级结构的典型特征，这 6 个 α 螺旋构成容纳小分子的疏水性内部腔，性信息素类似物呈弯曲状态进入其中（Sandler *et al*., 2000）。由于 AsCSP5 C 端有近 200 个氨基酸的延伸，使其具有 7 个 α 螺旋和 5 个 β 折叠结构，前面的 6 个 α 螺旋和直链串联结构与其他 7 个 AsCSPs 相似，其后具有 1 个 α 螺旋和 5 个 β 折叠结构（图 4：A）。这些二级结构的异同决定了这些基因的相似功能及其理化性质的多样性。

由于 AsCSP2 与 AgCSP2，以及 AsCSP7 与 AgCSP7 氨基酸序列一致性最高，分别为 95% 和 94%，所以将 AgCSP2 和 AgCSP7 作为同源建模模版，预测得到 AsCSP2 和 AsCSP7 的蛋白质三级结构（图4：B），结果也表明 AsCSP2 和 AsCSP7 都由多个螺旋、延长链及回折结构组成，其螺旋结构比较紧凑，界限分明，使整个蛋白质分子呈球状，构成 CSP 的各个 α 螺旋流动性很强，使 CSP 具备了可塑性的结构特点，目前在多种昆虫的 CSP 中已得到证实（Picimbon *et al*., 2000；Briand *et al*., 2002）。

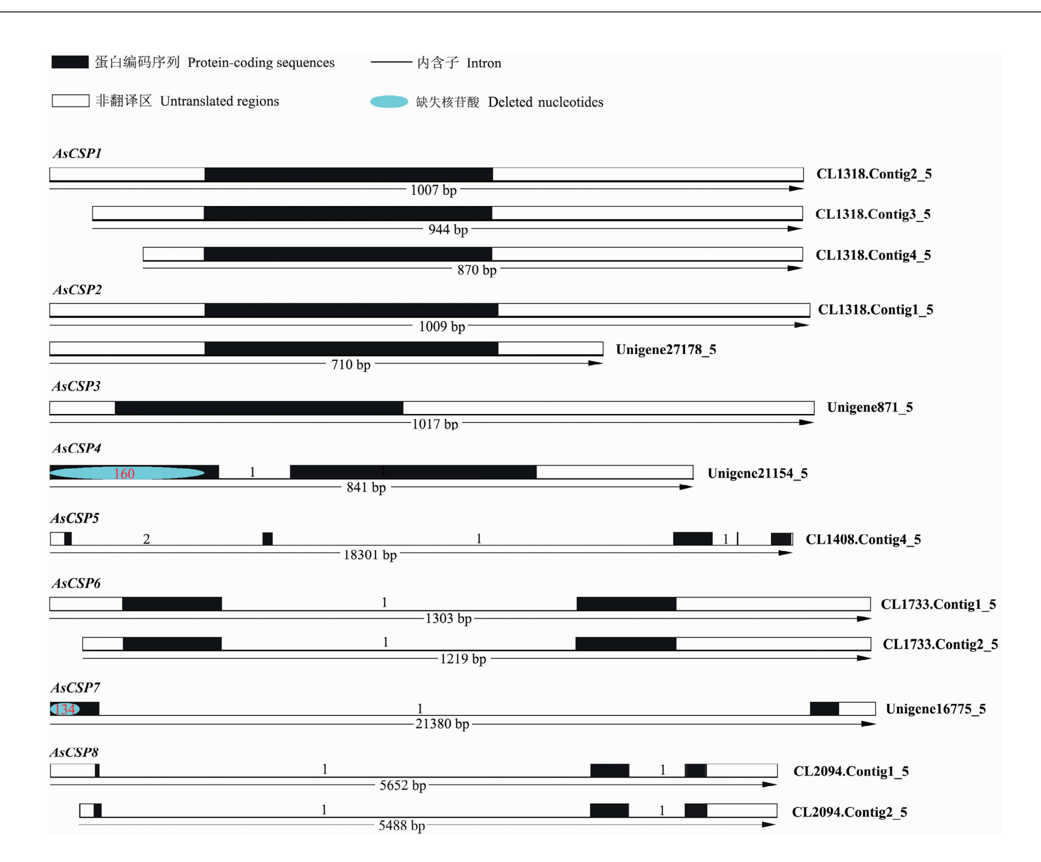


图 3 中华按蚊 CSP 家族基因内含子-外显子结构及选择性剪切

Fig. 3 The intron-exon structure and splicing variants of CSP genes from *Anopheles sinensis*

内含子相位用黑色数字标示在内含子横线的上方,红色数字标示此位点缺失核苷酸的长度,下方箭头和方向分别标示各转录本的长度(核苷酸数标于箭头线中部)和转录方向。Intron phases marked with black numerals above the horizontal lines demonstrate the introns, red numbers indicate the lengths of the deleted nucleotides at the positions, and arrows and directions below the lines denote the length and transcription direction of the transcripts (the nucleotide number is marked on the middle part of the arrow).

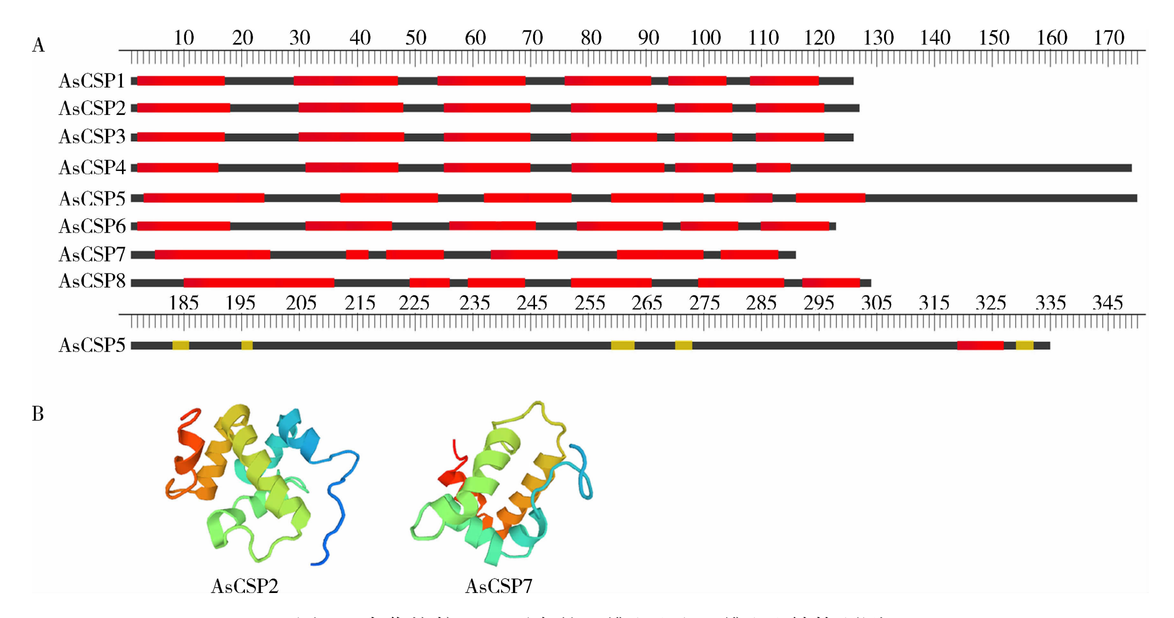


图 4 中华按蚊 CSP 蛋白的二维(A)和三维(B)结构预测

Fig. 4 Predicted 2D (A) and 3D (B) structure of CSPs from *Anopheles sinensis*

A: 二维结构 Two-dimensional structure. 每条序列的核苷酸位点用刻度尺标注,α 螺旋结构和 β 折叠结构分别以红色和黄色矩形标注。The amino acid positions of each sequence are marked by a ruler, and the α-helix structure and β-sheet structure are labelled with red and yellow rectangle, respectively. B: AsCSP2 和 AsCSP7 的三维结构图 Three-dimensional structures of AsCSP2 and AsCSP7.

2.5　双翅目代表性种 CSP 家族基因的系统发育关系

鉴定得到中华按蚊、冈比亚按蚊、埃及伊蚊和致倦库蚊分别有 8，8，43 和 27 个 CSP 家族基因，并更正了 *AgCSP5* 基因的序列。以黑腹果蝇 CSP 氨基酸序列为外群，利用 MEGA5.0 软件通过最大似然法对双翅目的 4 个蚊科种和 1 个果蝇科种的 CSP 家族基因构建了系统发育树（图 5）。结果显示，在 4 种蚊虫中都存在 1∶1 的直系同源基因，且形成较为分明的 8 个组（CSP1 – CSP8 group），都具有很好的自展值支持（bootstrap 为 51% ~100%）（图 5：A）。中华按蚊和冈比亚按蚊亲缘关系最近，*AgCSPs* 和 *AsCSPs* 的每个基因都分别对应地聚类在一个分支上，且其自展值为76% ~100%，这些基因在基因组

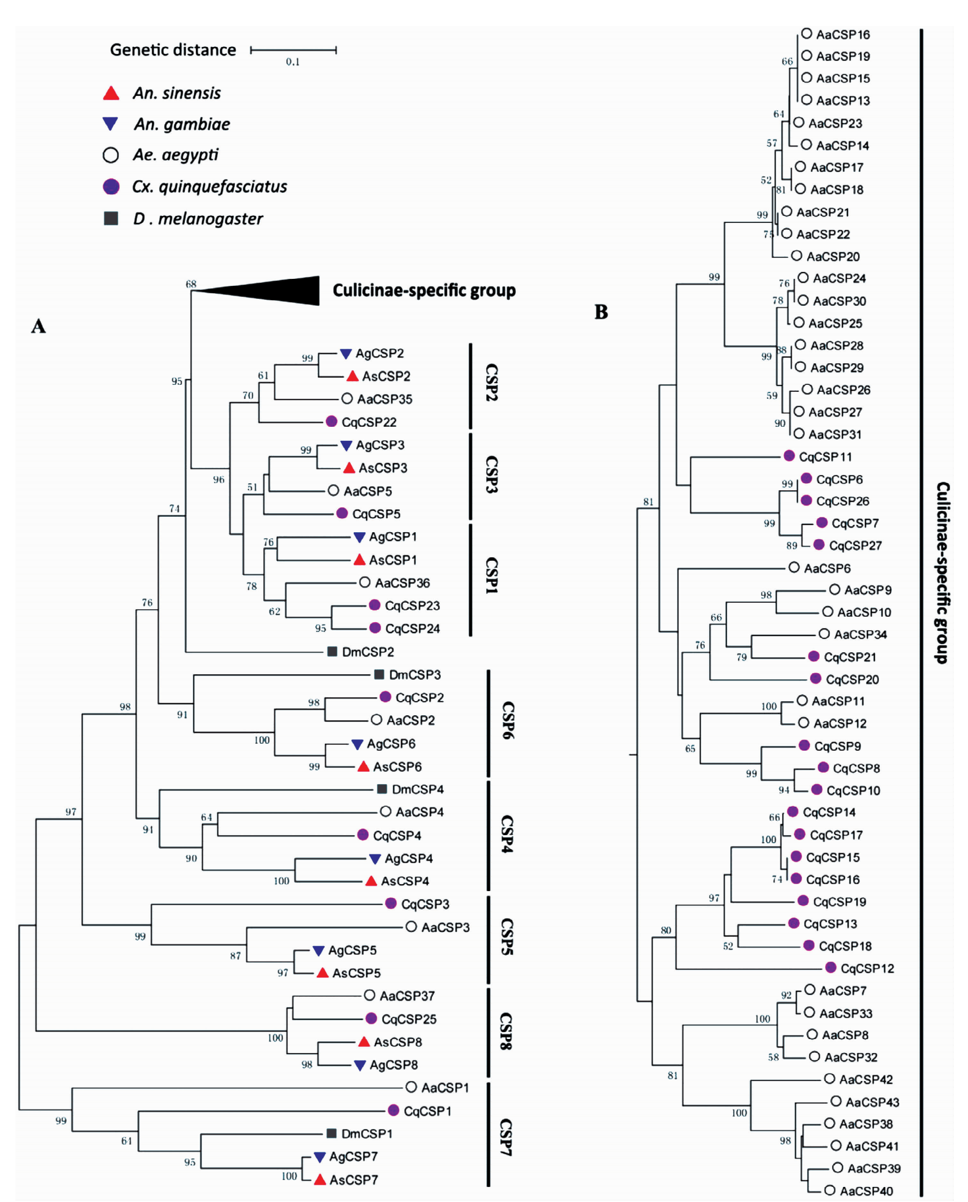


图 5　代表性双翅目昆虫基于氨基酸序列的 CSP 家族基因的系统发育关系（最大似然法）

Fig. 5　Phylogenetic relationships of the CSP gene family in representative Diptera insects based on the deduced amino acid sequences (maximum likelihood method)

As：中华按蚊 *Anopheles sinensis*；Ag：冈比亚按蚊 *Anopheles gambiae*；Aa：埃及伊蚊 *Aedes aegypti*；Cq：致倦库蚊 *Culex quinquefasciatus*；Dm：黑腹果蝇 *Drosophila melanogaster*. A：果蝇和 4 种蚊虫 CSP 家族基因的系统发育关系，其中库蚊亚科特异扩张的 Culicinae-specific 组以黑色三角标示出 Phylogenetic relationships of the CSPs of *D. melanogaster* and four mosquito species, and the Culicinae-specific group is shown as black triangle；B：库蚊亚科特异扩张的 Culicinae-specific 组内基因系统发育关系 Phylogenetic relationships of the genes in the Culicinae-specific group. 大于 50% 的 bootstrap 值标记在树的分支节点上；标尺代表系统遗传距离。Percentage bootstrap values higher than 50% are marked on each branch, and the scale bar indicates the genetic distance.

上的分布及系统发育分析表明 CSP 家族基因在这两个按蚊的进化中是相对保守的,彼此具有 1∶1 的直系同源关系,起源于这两个种的共同祖先。在 CSP4, CSP6 和 CSP7 组中,各有一个黑腹果蝇的 CSP 家族基因(*DmCSP*4, *DmCSP*3, *DmCSP*1),而 CSP5 和 CSP8 组中缺乏黑腹果蝇 CSP 同源基因。重要的是,黑腹果蝇的 *DmCSP*2 似乎和其他组的复合体(CSP1, CSP2, CSP3 和 Culicinae-specific group)形成了一个姊妹群。这个 Culicinae-specific 组的基因为库蚊亚科所特有(图 5: B),且具有 68% 的 bootstrap 值支持。在所研究的埃及伊蚊和致倦库蚊中的基因数远多于这两个种的其他 CSP 组(CSP1 - CSP8 group)基因的总和,分别有 35 个基因(总共 43 个基因)和 18 个基因(总共 27 个基因)(图 5: B)。致倦库蚊经历了一个基因的重复事件分化成 2 个 CSP1 基因(*CqCSP*23 和 *CqCSP*24)。

2.6 中华按蚊 CSP 家族基因的非同义替换

基因氨基酸编码区的核苷酸在进化过程中具有不同的替代速率,通常可利用非同义替换率与同义替换率的比值(Ka/Ks)来检验编码序列所经历的自然选择压力(周琦和王文, 2004)。我们选择中华按蚊与冈比亚按蚊的直系同源基因对计算其 Ka/Ks 比值,从数据结果可发现,这些基因对的 Ka/Ks 值都小于 1,介于 0.0131 ~ 0.2976 之间(表 3),说明 CSP 基因家族在进化过程中主要是受到纯化选择作用。其中,*CSP*4 和 *CSP*5 的 Ka/Ks 值相对较高,分别为 0.2267 和 0.2976,表明这两个基因所经受的选择压力更小。

表 3 中华按蚊与冈比亚按蚊 CSP 直系同源基因对的 Ka/Ks 分析

Table 3 Ka/Ks analysis for the orthologous CSP gene pairs between *Anopheles sinensis* and *An. gambiae*

CSP 直系同源基因对 Orthologous CSP gene pairs	Ka	Ks	Ka/Ks
*AsCSP*1-*AgCSP*1	0.0914	2.7835	0.0328
*AsCSP*2-*AgCSP*2	0.0229	0.6277	0.0364
*AsCSP*3-*AgCSP*3	0.0447	0.4616	0.0969
*AsCSP*4-*AgCSP*4	0.2245	0.9905	0.2267
*AsCSP*5-*AgCSP*5	0.2667	0.8961	0.2976
*AsCSP*6-*AgCSP*6	0.0389	2.9658	0.0131
*AsCSP*7-*AgCSP*7	0.0241	0.3917	0.0616
*AsCSP*8-*AgCSP*8	0.1193	1.2191	0.0979

Ka: 非同义替换率 Non-synonymous substitution rate; Ks: 同义替换率 Synonymous substitution rate. Ka/Ks > 1, 正选择 Positive selection; Ka/Ks = 1, 中性选择 Neutral selection; Ka/Ks < 1, 纯化选择 Purifying selection.

3 讨论

化学感受蛋白 CSP 在昆虫感受外界信息素的过程中扮演着重要的角色。这类蛋白在不同物种的数量差异很大,已知在冈比亚按蚊、黑腹果蝇、家蚕、赤拟谷盗、意大利蜜蜂、豌豆蚜和人虱分别有 8, 4, 19, 19, 6, 10 和 6 个 CSP 家族编码基因(Vieira and Rozas, 2011),表明不同适应性昆虫感受外界信息的分子机制存在相应差异。本研究在中华按蚊全基因组水平上鉴定得到 8 个 CSP 家族基因,中华按蚊气味分子结合蛋白 OBP 家族有 64 个基因(He *et al.*, 2016),进一步表明昆虫 CSP 基因数少于 OBP 基因数(Vieira and Rozas, 2011)。CSP 氨基酸序列比 OBP 在不同物种中更保守,CSP 比 OBP 的起源更早(Sanchez-Gracia *et al.*, 2009)。CSP 类似蛋白广泛存在于节肢动物中,但在脊椎动物中却极少(Vieira and Rozas, 2011),CSP 被认为是生物识别环境化学分子刺激的主要古老机制,随着自然选择压力的减小,高等动物中相应功能已逐渐被专化性更高的 OBP 和 PBP 家族基因所取代(徐浩智等, 2015)。在 OBP 存在的情况下,CSP 仍普遍存在于昆虫纲中,这表明 CSP 仍执行非常重要的生理功能。

中华按蚊的 8 个 AsCSPs 具有很高的保守性,其 OS-D 保守结构域为 85 或 93 个氨基酸,4 个半胱氨酸构成两个二硫键(CYS68-CYS75, CYS94-CYS97)。其蛋白结构分析发现,所有 AsCSPs 均有 6 个 α 螺旋,这个结构在其他物种中也普遍存在,也意味着这些基因的功能保守性。对转录组数据分析表明,这些 AsCSPs 基因均存在转录表达,*AsCSP*3 的表达量最高, FPKM 值达到 385.46。*AsCSP*2, *AsCSP*6 和 *AsCSP*8 具有两个选择性剪切,*AsCSP*1 具有 3 个选择性剪切。

本研究新鉴定了埃及伊蚊和致倦库蚊，再次鉴定了冈比亚按蚊全基因组上的 CSP 家族基因，和中华按蚊及黑腹果蝇一起研究了这些种的 CSP 家族基因的系统发育。由于4 种蚊虫中的8 个 CSP 基因都存在 1∶1 的直系同源基因并形成明显不同的支系，本研究把它们分为 8 个不同的组（CSP1 – CSP8 group）。在 CSP4，CSP6 和 CSP7 组中各存在一个黑腹果蝇的 CSP 基因，表明这些组的基因起源于蚊虫和果蝇分化之前，而 CSP5 和 CSP8 组中缺乏黑腹果蝇 CSP 同源基因，表明这两个组起源于蚊虫和果蝇分化之后。黑腹果蝇的 *DmCSP*2 和其他组的复合体（CSP1，CSP2，CSP3 和 Culicinae-specific group）是姊妹群，这个复合体的基因很可能是蚊虫在进化过程中为适应环境而大量扩增产生的。埃及伊蚊和致倦库蚊分别有 35 个基因和 18 个基因形成了一个特殊的支系，不为按蚊所共有，本研究将这个特有支系命名为 Culicinae-specific 组，这是库蚊亚科昆虫 CSP 家族基因的特异性扩张，反映了该亚科特有的进化和功能。这个 Culicinae-specific 组的埃及伊蚊基因数（35 个基因）明显多于致倦库蚊的基因数（18 个基因），这反映了库蚊亚科种间的差异。

自然界中发生的很多非同义突变都是有害突变，在净化选择的作用下这些位点的碱基替换率比较低。Ka/Ks 值越小，表明该基因承受的选择压力越大，保守程度越高。若 Ka/Ks 值 <1，表示同义替换的速率高于非同义替换的速率，否则认为有纯化选择的压力（周琦和王文，2004）。我们的研究表明所有中华按蚊与冈比亚按蚊同源基因对的 Ka/Ks 值都小于 1，这在一定程度上也说明昆虫基因组中 CSP 基因家族的进化相对缓慢且高度保守，从某种程度上也反映出这些化学感受蛋白在昆虫的生存中发挥着不可或缺的作用。

CSP 基因在昆虫触角、喙、下唇须和足等各化学感器中具有大量表达，担负着感受、识别和转运环境中化学刺激信息的功能。大量的研究也表明，这类蛋白在性腺、头部、腹部等没有任何化学感受结构的组织器官中表达（Jacquin-Joly *et al*.，2001），推断 CSP 在昆虫的生命活动中还参与许多其他重要的生理功能，如日本弓背蚁 *Camponotus japionicus* CSP 可溶解表皮碳氢类化合物，从而识别不同巢穴蚁群，并产生排异性（Ozaki *et al*.，2005）；蜜蜂和云杉色卷蛾 *Choristoneura fumiferana* CSP 在蜕皮前期的表达量明显高于蛹和成虫蜕皮期的表达量，被证实其与昆虫蜕皮有关（Wanner *et al*.，2005；Maleszka *et al*.，2007）；通过 RNA 干扰试验证实，甜菜夜蛾 *Spodoptera exigua* CSP 能显著降低其产卵和卵的孵化能力（Gong *et al*.，2012）。

本研究基于生物信息学分析在全基因组水平鉴定了中华按蚊 CSP 家族基因，分析了它们的特征，通过比较基因组分析讨论了这个家族基因在蚊科昆虫的进化，建立了新的分类系统，为后续开展 CSP 家族基因在感受外界信息及其他生理活动上的功能奠定了基础。

参考文献（References）

Andersson MN，Grosse-Wilde E，Keeling CI，Bengtsson JM，Yuen MM，Li M，Hillbur Y，Bohlmann J，Hansson BS，Schlyter F，2013. Antennal transcriptome analysis of the chemosensory gene families in the tree killing bark beetles，*Ips typographus* and *Dendroctonus ponderosae*（Coleoptera：Curculionidae：Scolytinae）. *BMC Genomics*，14(1)：1 – 16.

Angeli S，Ceron F，Scaloni A，Monti M，Monteforti G，Minnocci A，Petacchi R，Pelosi P，1999. Purification，structural characterization，cloning and immunocytochemical localization of chemoreception proteins from *Schistocerca gregaria*. *Eur. J. Biochem*.，262(3)：745 – 754.

Ban L，Zhang L，Yan Y，Pelosi P，2002. Binding properties of a locust's chemosensory protein. *Biochem. Bioph. Res. Commun*.，293(1)：50 – 54.

Briand L，Swasdipan N，Nespoulous C，Bézirard V，Blon F，Huet JC，Ebert P，Pernollet JC，2002. Characterization of a chemosensory protein（ASP3c）from honeybee（*Apis mellifera* L.）as a brood pheromone carrier. *Eur. J. Biochem*.，269(18)：4586 – 4596.

Chang XL，Zhong DB，Fang Q，Hartsel J，Zhou GF，Shi LN，Fang FJ，Zhu CL，Yan GY，2014. Multiple resistances and complex mechanisms of *Anopheles sinensis* mosquito：a major obstacle to mosquito-borne diseases control and elimination in China. *PLoS Negl. Trop. Dis*.，8(5)：e2889.

Chen B，Zhang YJ，He ZB，Li WS，Si FL，Tang Y，He QY，Qiao L，Yan ZT，Fu WB，Che YF，2014. *De novo* transcriptome sequencing and sequence analysis of the malaria vector *Anopheles sinensis*（Diptera：Culicidae）. *Parasit. Vectors*，7(1)：314.

Chow CY，1991. Malaria vectors in China. *Chin. J. Entomol. Special Publ*.，6(6)：67 – 79.

Dippel S，Oberhofer G，Kahnt J，Gerischer L，Opitz L，Schachtner J，Stanke M，Schütz S，Wimmer EA，Angeli S，2014. Tissue-specific transcriptomics，chromosomal localization，and phylogeny of chemosensory and odorant binding proteins from the red flour beetle *Tribolium castaneum* reveal subgroup specificities for olfaction or more general functions. *BMC Genomics*，15(1)：1 – 14.

Forêt S，Wanner KW，Maleszka R，2007. Chemosensory proteins in the honey bee：insights from the annotated genome，comparative analyses and expressional profiling. *Insect Biochem. Molec*.，37(1)：19 – 28.

Gong DP, Zhang HJ, Zhao P, Lin Y, Xia QY, Xiang ZH, 2007. Identification and expression pattern of the chemosensory protein gene family in the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.*, 37(3): 266–277.

Gong L, Luo Q, Rizwan-ul-Haq M, Hu MY, 2012. Cloning and characterization of three chemosensory proteins from *Spodoptera exigua* and effects of gene silencing on female survival and reproduction. *Bull. Entomol. Res.*, 102(05): 600–609.

Gu SH, Wang SY, Zhang XY, Ji P, Liu JT, Wang GR, Wu KM, Guo YY, Zhou JJ, Zhang YJ, 2012. Functional characterizations of chemosensory proteins of the alfalfa plant bug *Adelphocoris lineolatus* indicate their involvement in host recognition. *PLoS ONE*, 7(8): e42871.

Gu XC, Zhang YN, Kang K, Dong SL, Zhang LW, 2015. Antennal transcriptome analysis of odorant reception genes in the red turpentine beetle (RTB), *Dendroctonus valens*. *PLoS ONE*, 10(5): e0125159.

He X, He ZB, Zhang YJ, Zhou Y, Xian PJ, Qiao L, Chen B, 2016. Genome-wide identification and characterization of odorant-binding protein (OBP) genes in the malaria vector *Anopheles sinensis* (Diptera: Culicidae). *Insect Sci.*, 23(3): 366–376.

Hojo MK, Ishii K, Sakura M, Yamaguchi K, Shigenobu S, Ozaki M, 2015. Antennal RNA-sequencing analysis reveals evolutionary aspects of chemosensory proteins in the carpenter ant, *Camponotus japonicus*. *Sci. Rep.*, 5: 13541.

Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G, 2015. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, 31(8): 1296–1297.

Iovinella I, Bozza F, Caputo B, Della Torre A, Pelosi P, 2013. Ligand-binding study of *Anopheles gambiae* chemosensory proteins. *Chem. Senses*, 38(5): 409–419.

Jacquin-Joly E, Vogt RG, François MC, Nagnan-Le Meillour P, 2001. Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae*. *Chem. Senses*, 26(7): 833–844.

Jin X, Brandazza A, Navarrini A, Ban L, Zhang S, Steinbrecht RA, Zhang L, Pelosi P, 2005. Expression and immunolocalisation of odorant-binding and chemosensory proteins in locusts. *Cell. Mol. Life Sci.*, 62(10): 1156–1166.

Lartigue A, Campanacci V, Roussel A, Larsson AM, Jones TA, Tegoni M, Cambillau C, 2002. X-ray structure and ligand binding study of a moth chemosensory protein. *J. Biol. Chem.*, 277(35): 32094–32098.

Li Z, Shen Z, Zhou J, Field L, 2003. Bioinformatics-based identification of chemosensory proteins in African malaria mosquito, *Anopheles gambiae*. *Genomics, Proteomics & Bioinformatics*, 1(4): 288–298.

Li ZQ, Zhang S, Luo JY, Zhu J, Cui JJ, Dong SL, 2015. Expression analysis and binding assays in the chemosensory protein gene family indicate multiple roles in *Helicoverpa armigera*. *J. Chem. Ecol.*, 41(5): 473–485.

Liu BQ, Qiao L, Xu BY, Zheng XL, Chen B, 2016. Identification and characterization of the CPF family of cuticular protein genes in the genome of *Anopheles sinensis* (Diptera: Culicidae). *Acta Entomologica Sinica*, 59(6): 622–631. [刘柏琦, 乔梁, 许柏英, 郑学令, 陈斌, 2016. 中华按蚊 CPF 家族表皮蛋白基因的全基因组鉴定及其特征分析. 昆虫学报, 59(6): 622–631]

Maleszka J, Foret S, Saint R, Maleszka R, 2007. RNAi-induced phenotypes suggest a novel role for a chemosensory protein csp5 in the development of embryonic integument in the honeybee (*Apis mellifera*). *Dev. Genes Evol.*, 217(3): 189–196.

Manguin S, Bangs M, Pothikasikorn J, Chareonviriyaphap T, 2010. Review on global co-transmission of human *Plasmodium* species and *Wuchereria bancrofti* by *Anopheles* mosquitoes. *Infect. Genet. Evol.*, 10(2): 159–177.

McKenna MP, Hekmat-Scafe DS, Gaines P, Carlson JR, 1994. Putative *Drosophila* pheromone-binding proteins expressed in a subregion of the olfactory system. *J. Biol. Chem.*, 269(23): 16340–16347.

Monteforti G, Angeli S, Petacchi R, Minnocci A, 2002. Ultrastructural characterization of antennal sensilla and immunocytochemical localization of a chemosensory protein in *Carausius morosus* Brünner (Phasmida: Phasmatidae). *Arthropod Struct. Dev.*, 30(3): 195–205.

Ozaki M, Wada-Katsumata A, Fujikawa K, Iwasaki M, Yokohari F, Satoji Y, Nisimura T, Yamaoka R, 2005. Ant nestmate and non-nestmate discrimination by a chemosensory sensillum. *Science*, 309(5732): 311–314.

Pelletier J, Leal WS, 2011. Characterization of olfactory genes in the antennae of the southern house mosquito, *Culex quinquefasciatus*. *J. Insect Physiol.*, 57(7): 915–929.

Pelosi P, Calvello M, Ban L, 2005. Diversity of odorant-binding proteins and chemosensory proteins in insects. *Chem. Senses*, 30(Suppl. 1): i291–i292.

Pelosi P, Zhou JJ, Ban LP, Calvello M, 2006. Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci.*, 63(14): 1658–1676.

Picimbon JF, Dietrich K, Breer H, Krieger J, 2000. Chemosensory proteins of *Locusta migratoria* (Orthoptera: Acrididae). *Insect Biochem. Molec. Biol.*, 30(3): 233–241.

Pikielny CW, Hasan G, Rouyer F, Rosbash M, 1994. Members of a family of *Drosophila* putative odorant-binding proteins are expressed in different subsets of olfactory hairs. *Neuron*, 12(1): 35–49.

Qin Z, Ran YH, Zhi ZJ, Yan ZT, Zhang YJ, Huang T, He ZB, Chen B, 2014. Cloning and expression analysis of an odorant binding protein gene *AsinOBP*1 from *Anopheles sinensis* (Diptera: Culicidae). *Acta Entomologica Sinica*, 57(11): 1289–1298. [秦赠, 冉永红, 支中婧, 闫振天, 张玉娟, 黄婷, 何正波, 陈斌, 2014. 中华按蚊气味结合蛋白基因 *AsinOBP*1 的克隆和表达分析. 昆虫学报, 57(11): 1289–1298]

Robertson HM, Martos R, Sears CR, Todres EZ, Walden KK, Nardi JB, 1999. Diversity of odourant binding proteins revealed by an expressed sequence tag project on male *Manduca sexta* moth antennae. *Insect Mol. Biol.*, 8(4): 501–518.

Sabatier L, Jouanguy E, Dostert C, Zachary D, Dimarcq JL, Bulet P,

Imler JL, 2003. Pherokine-2 and -3. *Eur. J. Biochem.*, 270(16): 3398 – 3407.

Sanchez-Gracia A, Vieira FG, Rozas J, 2009. Molecular evolution of the major chemosensory gene families in insects. *Heredity*, 103(3): 208 – 216.

Sandler BH, Nikonova L, Leal WS, Clardy J, 2000. Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. *Chem. Biol.*, 7(2): 143 – 151.

Schneider D, 1969. Insect olfaction: deciphering system for chemical messages. *Science*, 163(3871): 1031 – 1037.

Tang Y, Qiao L, Zhang YJ, Che YF, Hong R, Chen B, 2014. Identification and bioinformatics analysis of genes of the CYP6Y subfamily in *Anopheles sinensis* (Diptera: Culicidae). *Acta Entomologica Sinica*, 57(6): 663 – 672. [唐尧, 乔梁, 张玉娟, 车艳飞, 洪瑞, 陈斌, 2014. 中华按蚊 CYP6Y 亚家族基因的鉴定和生物信息学分析. 昆虫学报, 57(6): 663 – 672]

Vieira FG, Rozas J, 2011. Comparative genomics of the odorant-binding and chemosensory protein gene families across the Arthropoda: origin and evolutionary history of the chemosensory system. *Genome Biol. Evol.*, 3: 476 – 490.

Wang XY, Xiong M, Lei CL, Zhu F, 2015. The developmental transcriptome of the synanthropic fly chrysomya megacephala and insights into olfactory proteins. *BMC Genomics*, 16: 20.

Wanner KW, Isman MB, Feng Q, Plettner E, Theilmann DA, 2005. Developmental expression patterns of four chemosensory protein genes from the eastern spruce budworm, *Chroistoneura fumiferana*. *Insect Mol. Biol.*, 14(3): 289 – 300.

Wanner KW, Willis LG, Theilmann DA, Isman MB, Feng Q, Plettner E, 2004. Analysis of the insect OS-D-like gene family. *J. Chem. Ecol.*, 30(5): 889 – 911.

Xu HZ, You YW, Zhang L, 2015. A review on chemosensory protein and its functions in insects. *Journal of Agricultural Biotechnology*, 23(1): 118 – 125. [徐浩智, 游银伟, 张龙, 2015. 昆虫化学感受蛋白及其功能研究进展. 农业生物技术学报, 23(1): 118 – 125]

Yang Z, Nielsen R, 2000. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.*, 17(1): 32 – 43.

Zhang SQ, Cheng F, Webber R, 1994. A successful control programme for lymphatic filariasis in Hubei, China. *Trans. R. Soc. Trop. Med. Hyg.*, 88(5): 510 – 512.

Zhou Q, Wang W, 2004. Detecting natural selection at the DNA level. *Zoological Research*, 25(1): 73 – 80. [周琦, 王文, 2004. DNA 水平自然选择作用的检测. 动物学研究, 25(1): 73 – 80]

（责任编辑：马丽萍）